Current Practices in Fine Art Reproduction: Project Summary

Susan Farnand ***and Jun Jiang, Rochester Institute of Technology, Rochester, NY and Franziska Frey, Harvard College Library, Cambridge, MA***

Abstract

Many cultural heritage institutions are currently spending significant resources photographing their works of art for a variety of applications with distinctly different requirements. To create reproductions of their artwork, cultural heritage institutions employ a range of technology and a variety of workflows. A similar variety is used to publish these images in a number of output media. This project was undertaken to explore these workflows and the image quality of the reproductions they generate. The objectives of this project were to: (1) determine the optimal reproduction processes in use in cultural heritage institutions today, (2) document the image quality inherent in current workflows in print and online, (3) define key quality criteria based on objective and subjective metrics, and (4) use this information to develop a framework to serve as a guideline for museums to follow when reproducing fine art. To work towards these objectives, a series of experiments were developed to evaluate the image quality attainable with the current reproduction workflows.

Key findings of the project included that (1) achieving accurate tone reproduction at capture is crucial, (2) acceptable reproductions are achievable using a digital press, (3) following standardized workflows, ISO printing standards, and viewing standards substantially reduces the need for manual post-processing, (4) camera make, lighting, and file format had little impact on the ranking results, (5) internet-based experiments may be successfully used when evaluating image preference and, (6) while workflows still vary considerably, some commonalities were found for workflows producing images that were generally ranked highly across the experiments. These workflows were used as a basis for the development of the recommended guidelines, which included the following recommendations and considerations:

- *Workflows covering the whole image interchange cycle should be documented in detail. No undocumented processing should be performed along the image interchange cycle.*
- *ICC profile-based color management should be used to achieve best results.*
- *The use of targets to ensure a proper capture setup is recommended.*
- *Defining imaging goals and talking to users is indispensable to help set expectations.*
- *Guide prints did not prove useful in these experiments and are not recommended as proofs (though more testing may be needed). They could, however, be used for a visual 'reality check' on press*
- *Closing the communication loop in the image interchange cycle is of the utmost importance.*

The three-year project was financially supported by The Andrew W. Mellon Foundation. Detailed information on this project, including the final report, can be found at www.artimaging.rit.edu.[1,2]

Introduction

A series of experiments were conducted to evaluate the image quality attainable with the current reproduction workflows. The experiments were developed in order to benchmark art image interchange cycles for printed (both lithographic and digital) and on-screen reproductions. The experiments were conducted with and without the original present. Seventeen institutions captured a variety of artwork and objective targets. The objective targets were included with the goal of finding measurable image characteristics that correlated with the subjective results. Based on the experimental results, along with interviews conducted with a range of participants in Image Interchange cycles, a set of guidelines was developed.

Experimental

General methodology

The artwork used in this study included six pictorial works intended to provide a variety of images for exploring the impact of scene content and various image media while keeping the number of test stimuli manageable. All artwork was commissioned or acquired from sources that allowed for unrestricted reproductions rights. The pieces included an aquatint entitled '*Still Life'*, a historic platinum print photographic portrait ('*Photograph'*), an acrylic painting entitled '*Firelight'*, a watercolor entitled '*Mountain'*, and two oil paintings: '*Daisies'*, and 'Inspired by Monet's Waterloo Bridge' ('*Bridge*'). The artwork is shown in Figure 1.

Figure 1*. The artwork for the Current Practices in Fine Art Reproduction project, from top left: Photograph, Daisies (oil), Mountain (watercolor), Bridge (oil), Still Life (aquatint), and Firelight (acrylic)*

It was hypothesized that color and tone reproduction, sharpness, and uniformity might be among the key image characteristics in determining the perceptual quality of fine art reproductions. As such, the objective targets used in the experiment included: a Macbeth Color Checker and a paint patch target for assessing color and tone measurement throughout the reproduction process, and the Universal Test Target® from Image Engineering for assessing color, tone, sharpness, and uniformity at capture, Figure 2.

The familiar Macbeth Color Checker is a target that was created in the 1970s specifically for the evaluation of photographic color reproduction.[3] The paint target was constructed for this study using the same paints used to make the oil paint pictorial targets. And the Universal Test Target (UTT) is a commercially available target developed to evaluate the quality of the image capture. (Other commercially available test targets used for assessing image quality at capture in fine art reproduction include the Golden Thread® target from Image Science Associates.)

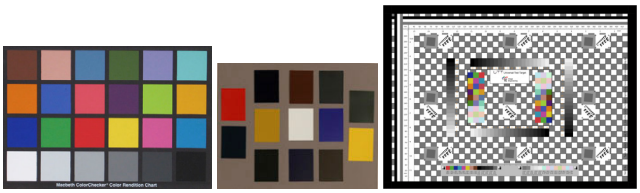

***Figure 2**. The targets for the Current Practices in Fine Art Reproduction project, from left: Macbeth Color Checker, paint patch target, Universal Test Target®*

Each institution delivered digital files and information regarding its workflow. If the institution normally would provide guide prints to their printers, then they were directed to supply them for this study. Prints from the digital files were then generated on the Heidelberg Speedmaster at the RIT Printing Applications Laboratory (PAL) using Prinergy Workflow 3.0.2.2® and Kodak Thermal Gold® plates on a VLF 5080 Quantum® plate setter and following the ISO 12467 protocol.

Using the prints made from the delivered digital files, a series of psychophysical experiments were conducted to generate relative visual ratings of image quality. These included a study of the impact of viewing lighting on the perception of image quality, an evaluation of the fine art reproduction workflow on perceived image quality both in print and on an LCD display, an evaluation of the effect of the presence of the original on perceived image quality in print and on an LCD display, and a comparison of the perceived image quality of offset lithographic prints to electrophotographic prints. The list of the experiments conducted and the objectives of each is delineated in Table I.

The experiments primarily took place in the Display and Perception Lab in the Color Science Hall at RIT. The prints and the originals were viewed under D50 lighting conditions. The experimental setups are shown in Figures 3.

***Figure 3.** The hardcopy (left) and softcopy (right) experimental setups*

The observers who participated in the experimentation for this project generally consisted of individuals in the cultural heritage community. The observers included librarians, photographers, curators, art teachers, conservators, and imaging science students and staff with an interest in fine art reproduction. Observers were tested for color vision anomalies with Ishihara plates.

Table 1. Experiments conducted as part of the Current Practices in Fine Art Reproduction project

Experiment Title	Objective
The Impact of Lighting on Perceived Quality of Fine Art Reproductions[4]	Evaluate the effect of changes in viewing illumination on the perceived quality of printed fine art reproductions
Evaluating CATs as Predictors of Observer Adjustments in Softcopy Fine Art Reproduction[5]	Evaluate trends in color adjustments made by museum, library, and archive personnel to artwork reproductions presented on displays in order to better understand how images should be processed to effectively represent original artwork on screen
The Impact of Workflow on Perceived Reproduction Quality of Fine Art Images[6,7]	Understand the relative reproduction quality being achieved by fine art workflows in use today
The Impact of Workflow on Perceived Quality of Fine Art Images[6]	Understand the relative perceived quality being achieved by fine art workflows in use today in generating images being evaluated without the original artwork present
The Perceived Image Quality Of Printed Fine Art Reproductions[8]	Compare the relative perceived image quality of prints made on an offset lithographic press to those made on an electrophotographic digital press
The Impact of Workflow on Perceived Quality of Fine Art Images being Viewed on the Web[9]	Understand the relative quality being achieved by fine art workflows in use today in generating images to be viewed on the web

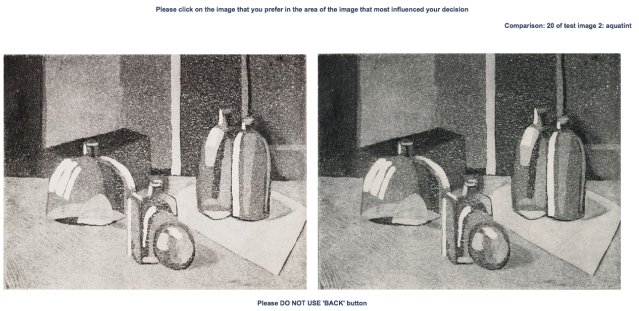

Figure 4. *Graphic user interface for web-based ranking experiment*

The Impact of Workflow on Perceived Quality of Fine Art Images being Viewed on the Web

In the internet-based experiment, a web application was implemented to evaluate the impact of workflow on perceived quality of fine art reproductions in this uncontrolled environment. Observers from around the world were invited to take part in the experimentation, with limited constraints on viewing conditions as long as reasonable Internet speed and a web browser were available. The main interface of the web-based application is shown in Figure 4. In the experiment, observers were asked to select their preferred image for each in a series of image pairs; the same protocol as used in the lab experiment.

In order to examine how observers' display settings affected the ranking results, observers were requested to answer questions prior to the ranking experiment designed to provide information regarding their display contrast and color management. Observers were asked to identify the number of shades visible in two grayscales, one having a continuum of neutral shades and the second being composed of two black, three gray, and two whitish shades. The latter was intended to test the performance of the display in showing just-noticeable-differences (JND) in highlight, mid-tone and dark neutral colors. The number of shades that were identified by observers would be indicative of the gamma configuration (and thus the contrast) of the display. A vast majority of observers were able to distinguish all twelve shades in the top grayscale. However, observers were not always able to see the differences in the second grayscale. About half of the observers were unable to distinguish all seven shades in the JND-grayscale, indicating a too-low or too-high contrast resulting from an improper setting of gamma on the display.

In addition to the question regarding display contrast, observers' web browser color management modules were also of interest. A web browser with effective color management settings can identify ICC profiles when they are embedded in an image, and thus are able to display images in the intended colors. To determine whether the browser being used could successfully identify embedded profiles, the participants were asked to simultaneously view two images that were tagged with sRGB and AdobeRGB ICC profiles.. If they could distinguish differences in color between the two images, then color management must have been supported by the web browser. Since the images provided by the sixteen institutions were not tagged with the same ICC profiles, color management support would be beneficial to the evaluation of perceived image quality in the web-based ranking experiment. At least one web browser did not employ effective color management.

Only nine of the 95 observers used browsers that appeared to not be color managed. Since the observers using color-managed browsers outnumbered those who did not by a large margin, statistical analysis on the influence of web browsers was not possible.

Results

The Impact of Lighting on Perceived Quality of Fine Art Reproductions

Art reproductions are viewed under various lighting conditions, from the gallery to the museum shop to the living room. It is important, consequently, to understand and quantify the effect of lighting changes. In this experiment, the print reproductions were ranked under D50 and 'Horizon' lighting conditions. The results indicate that lighting conditions had a significant impact on the relative rankings, especially for the prints of the historic photograph. For this image, the best renditions under D50 lighting were the worst under Horizon lighting and vice versa. This suggests that prints made to match under one lighting specification may look 'just wrong' under another. Also, far fewer prints were considered acceptable reproductions of the original under Horizon lighting, which is not surprising since the digital files were created assuming a D50 workflow.

The Impact of Workflow on Perceived Reproduction Quality of Fine Art Images

The Impact of Workflow on Perceived Quality of Fine Art Images

Understanding which workflows provide both acceptable representations of the originals as well as pleasing images is important for those who are striving to provide images of fine art. In these experiments, which constituted the main body of work in the project, images generated from the nineteen workflows were evaluated in print and on electronic displays both with the original present (Perceived Reproduction Quality) and without (Perceived Quality). The results indicated that there were limited differences between displayed image reproductions relative to printed ones. Further, the differences between rankings made with and without the original artwork present were much smaller than expected, especially for the printed images. For the prints, the most significant changes in the image rankings occurred for the prints generated from the more monochromatic originals. These originals were not neutral, but, when they were not available for comparison, people preferred the more neutral renditions.

For the soft-copy mode, the differences in the image rankings with and without the original present were more significant. The workflow most often selected as providing the best representation of the original was considered the worst, on average, when the original was not present. The tone curve, as measured at capture, for this workflow is plotted in Figure 5. Note that the darkest inputs are significantly lighter. This absence of black was not something the participants in the experiment found appealing. Generally, when the original was not present, meaning the participants were ranking the images by preference, the lower contrast renditions were ranked lower and higher contrast ones higher. The tone curve shown in Figure 6 is was captured for a workflow that did not rank well when the original was present and did substantially better when it was not. These results may indicate

that files being prepared for the web may need to be processed differently from those being prepared for print.

A few workflows included in the study achieved relatively highly-ranked results for print and display, with and without the original. The tone curve for one of these workflows, which did particularly well for the chromatic originals in the study, is shown in Figure 7. This workflow, along with three others that generally achieved higher than average rankings, involved only minimal post-processing.

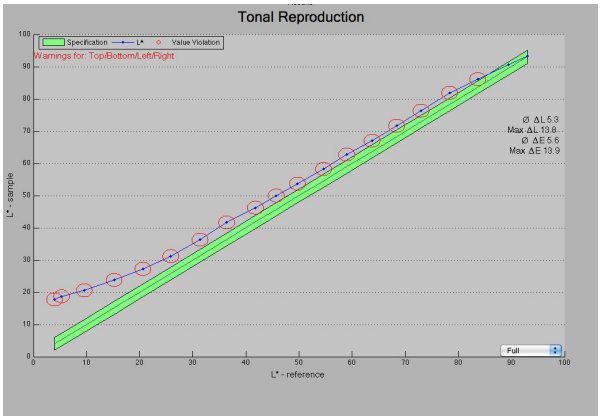

Figure 5. Tone reproduction curve for the UTT for the workflow that went from top-ranked for color reproduction to the lowest ranked for preference

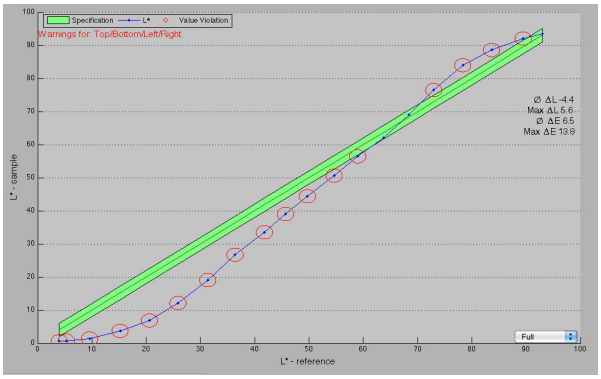

Figure 6. Tone reproduction curve for the UTT for the workflow that was ranked poorly for color reproduction but was among the more preferred images for several pieces of artwork

Eleven of the participating institutions provided guide prints to use when the prints were made for the hard copy experiments. The addition of these guide prints did not significantly increase (for one workflow, they significantly decreased) the rankings for their corresponding workflows, on average. It must be stated, however, that only limited adjustments could be made on press; new plates were not made in this study if adjustments on press could not generate a match.

The Perceived Image Quality of Printed Fine Art Reproductions

To supplement the visitors' cultural heritage experience, institutions print materials such as posters and cards. Generally, they provide materials only for their most popular holdings due to the high cost of setup for lithographic presses. Avoiding these costs by using digital presses would open up opportunities to print material for a wider range of artwork, possibly on-demand. To explore the image quality differences that might result, prints from the delivered files of the eight participating institutions that currently generate the most printed material were generated on a liquid electrophotographic digital press as well as a Heidelberg Speedmaster sheetfed lithographic press and used as stimuli in psychometric testing in which observers were asked to evaluate the prints as reproductions of the original artwork. The results indicated that there were limited differences in perceived image quality of printed reproductions made using offset and digital technologies when the transformation of the RGB files into CMYK images was performed following the same standard procedure (ISO 12647).

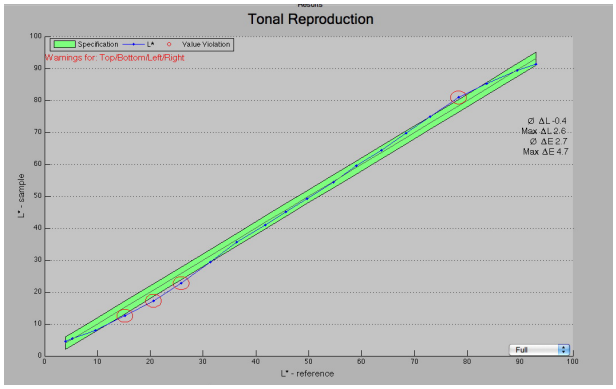

Figure 7. Tone reproduction curve for the UTT for the workflow that was ranked a little above average for color reproduction and among the best for preference

The Impact of Workflow on Perceived Quality of Fine Art Images being Viewed on the Web

Ninety-five observers, ranging in age from 20 to 65, participated in the experiment. There were approximately an equal number of male and female participants. The size of the displays used by observers ranged from 13 inches to 30 inches. About two-thirds of participants used the Macintosh® operating system while the rest used the Microsoft Windows® platform. The vast majority of observers took part in the experiment through either the Firefox® or Safari® web browsers, with a few using Internet Explorer® or Google Chrome®.

The results of this experiment showed a strong correlation between the rankings in the web-based experiment and those from the soft copy experiment without the original present. The Spearman rho correlations between the web-based ranking results and those obtained in the lab environment without the original present were highly significant for all six originals,. As an example, the z-scores converted from the ranking data in all three experiments for *Firelight* are shown in Figure 5. This result indicates that experimental conditions had minimal impact on preference judgments of images.

The experimental results from observers who could not identify all seven shades in the JND-grayscale did not have significantly different results from those of observers who

identified all seven shades despite the differences in the achievable contrast on observers' displays. In other words, the perceived image quality—when evaluated on the basis of preference—was not significantly affected by the contrast settings of the display. These results are important indicators that, when determining preference rankings, internet-based experiments may prove effective.

In Figure 8, almost all of the data points aligned along the regression line in the plot on the right, indicating a high correlation between the rankings of images in the web-based experiment and those obtained without the original present in the lab environment. However, in the plot on the left, the data points were scattered, indicating a lack of correlation between the ranking results in the web experiment and those obtained with the original present in the lab environment.

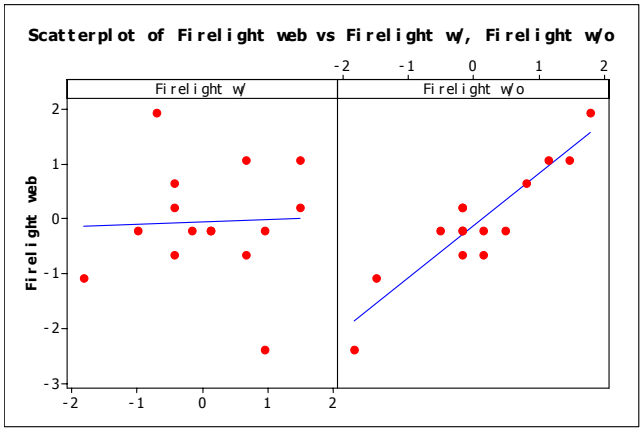

Figure 8. *Z-scores for Firelight rankings from the web-based experiment versus the soft copy results with and without the original present*

A significantly high correlation was found between the ranking results in the web-based experiment and the results from the soft copy experiment without the original present. Given the high correlation and the absence of the originals in both experiments, a similar criterion—image preference—must be shared by observers when the images were evaluated. Moreover, the preference judgments by observers were less likely to be affected by display settings, lighting conditions, and so forth, given the wide range of displays used by observers who participated in the web-based experiment. One direct application of this result would be the ability to conduct evaluations of image quality based on preference online when a characterized display or a controlled environment was not readily available. Similar results could be expected when having observers perform the test through the Internet as in the lab environment. However, a relatively large number of observers would be needed to eliminate the biases introduced by various testing conditions.

Conclusion

Key findings of the project included:

- The experimental results made it possible to develop a method to connect objective, measurable image quality to subjective image quality as perceived by observers.
- Workflows still vary considerably, but some commonalities were found for workflows producing images that were generally ranked highly across the experiments. These workflows were used as a basis for the development of the recommended guidelines.
- Workflows covering the whole image interchange cycle should be documented in detail. No undocumented processing should be performed along the image interchange cycle.
- Camera make, lighting, and file format had little impact on the ranking results
- The workflows that most accurately captured the toner curve, as measured using the Universal Test Target®, generally achieved the highest rankings and did not require extensive manual processing
- Guide prints did not prove useful in these experiments
- ICC profile-based color management should be used to achieve best results.
- The use of targets to ensure a proper capture setup is recommended.
- Following standardized workflows, ISO printing standards, and viewing standards substantially reduces the need for manual post-processing.
- Manual post-processing could not entirely overcome a poor capture
- Lighting conditions have a strong impact on image appearance.
- Acceptable reproductions are achievable using a digital press.
- Defining imaging goals and talking to users is indispensable to help set expectations.
- Closing the communication loop in the image interchange cycle is of the utmost importance.
- Workflow efficacy was determined to be the same whether reproductions were viewed in print or on electronic display.
- Internet-based experiments may be successfully used when evaluating image preference.

References

[1] McCamy, C. S., Marcus, H., and Davidson, J. G., "*A Color-Rendition Chart*", Journal of Applied Photographic Engineering, 2, 95-99 (1976).

[2] Farnand, S. P., Frey, F .S., The Impact of Lighting on Perceived Quality of Fine Art Reproductions, Proceedings of AIC's Conference on the Interaction of Colour and Light, Zurich, Switerland, June 2011.

[3] Farnand, S. P., Frey, F. S., Perceived Image Quality of Printed Fine Art Reproductions Proceedings of IS&T's Non-Impact Printing Conference, Minneapolis, MN, October 2011.

[4] Farnand, S. P., Jiang, J., Frey, F .S., Comparing hardcopy and softcopy results in the study of the impact of workflow on perceived reproduction quality of fine art images, Proceedings of the Image Quality and Systems Performance VII conference at the IS&T/SPIE Electronic Imaging Symposium, San Francisco, CA, January 2011.

[5] Frey, F .S., Farnand, S. P., Benchmarking Art Image Interchange Cycles, Final Report for The Andrew W. Mellon Foundation, www.artimaging.rit.edu, June 2011.

[6] Farnand, S. P., Jiang, J., Frey, F .S., Investigation into the impact of tone reproduction on the perceived image quality of fine art reproductions, Proceedings of the Image Quality and Systems

Performance VII conference at the IS&T/SPIE Electronic Imaging Symposium, San Francisco, CA, January 2011.

[7] Jiang, J., Frey, F .S., Farnand, S. P., Evaluating CATs as Predictors of Observer Adjustments in Softcopy Fine Art Reproduction, Proceedings of IS&T's Color Imaging Conference, San Antonio, TX, November 2010.

[8] Jiang, J., Frey, F .S., Farnand, S. P., Evaluating the Perceived Quality of Soft-copy Reproductions of Fine Art Images With and Without the Original Present, Proceedings of IS&T's Color Imaging Conference, San Jose, CA November 2011.

Author Biography

Susan Farnand received her BS in engineering from Cornell University and her Masters in Imaging Science from the Rochester Institute of Technology. After beginning her career at Eastman Kodak, she moved to RIT where she works as a Research Scientist in the Center for Imaging Science. Her research interests include image quality, human vision and perception, and color science. She is a member of IS&T and serves as an Associate Editor for the Journal of Imaging Science and Technology, and has served as co-chair of the IQSP conference at EI and guest editor for the Journal of Electronic Imaging.

.